CSE 559A: Computer Vision

[credit: danjodon.deviantart.com]

Fall 2017: T-R: 11:30-1pm @ Lopata 101

Instructor: Ayan Chakrabarti (ayan@wustl.edu).
Staff: Abby Stylianou (abby@wustl.edu), Jarett Gross (jarett@wustl.edu)

http://www.cse.wustl.edu/~ayan/courses/cse559a/

Oct 3, 2017

GENERAL

- PSET 1 Prob 6 Solutions posted for use in PSET 2.

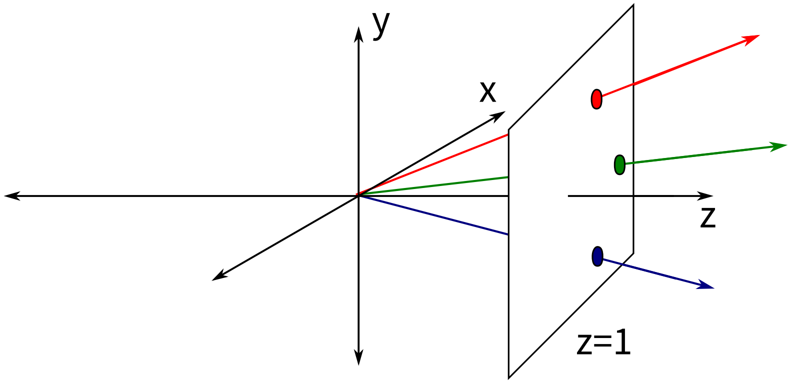
- Typo in PSET 2 code: ntod should return HxW array (not HxWx3)

- Still issues with monday office hours location
  - Tentatively still at Jolley 431
  - Keep an eye out on Piazza
  - [If J431 is locked, check collaboration area outside J517]

- Recitation this Friday Oct 6.

HOMOGENEOUS CO-ORDINATES

- Useful way to think about 2-D Homogeneous Co-ordinates $\mathbb{P}^2$



"Rays" in $\mathbb{R}^3$

- Cartesian form is "intersection" with plane $z = 1$.
- $(x, y, 0)$ are forms that are parallel to the $z = 1$ plane, intersect at infinity.
- 3-D Homogeneous Co-ordinates are rays in 4D, intersection with a hyper-plane.

QUICK WORD ABOUT NOTATION

- We assume vectors are column vectors.
- $p' = [x, y]$ implies a 2-D row vector (of size 1×2)
- $p = [x, y]^T = \begin{bmatrix} x \\ y \end{bmatrix}$ implies a 2-D column vector (of size 2×1)

HOMOGENEOUS CO-ORDINATES: 2D

Lines

Equation of a line in 2D:

$$ax + by + c = 0$$

Let $p = [\alpha\, x, \alpha\, y, \alpha]^T$ be homogeneous co-ordinates of a point (x, y). Then,

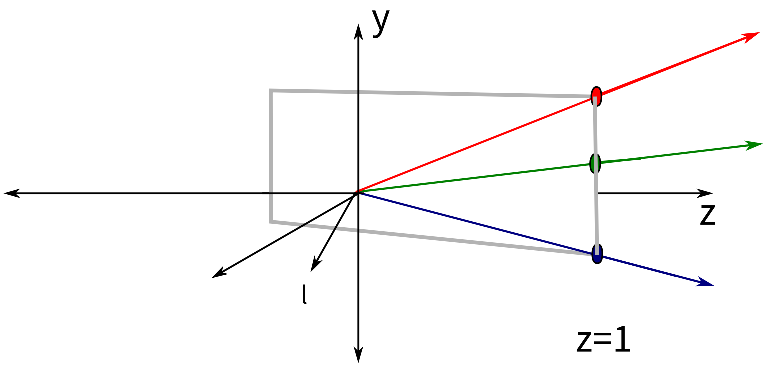
$$l^T p = 0, \quad l = [a, b, c]^T$$

Interestingly, l is also defined "upto scale": $l' = [\beta a, \beta b, \beta c]^T$ describes the same line as l.

HOMOGENEOUS CO-ORDINATES: 2D

Lines

Since $l^T p = 0$ for all points that lie on a line:



All rays for the line lie on a plane

l is the normal vector to that plane

HOMOGENEOUS CO-ORDINATES: 2D

Lines

Given two points p_1 and p_2, what is the homogeneous vector for the line joining them ?

It has to be an l such that $l^T p_1 = 0$ and $l^T p_2 = 0$.

Is that sufficient to determine l ?

Yes. Because, only need l upto scale.

Solution given by: $l = p_1 \times p_2$ (**Vector Cross-product**)

Recap: Writing $u = [u_1, u_2, u_3]^T = u_1 \hat{i} + u_2 \hat{j} + u_3 \hat{k}$, and $u = [v_1, v_2, v_3]^T = v_1 \hat{i} + v_2 \hat{j} + v_3 \hat{k}$

$$u \times v = \det \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = (u_2 v_3 - u_3 v_2)\hat{i} + (u_3 v_1 - u_1 v_3)\hat{j} + (u_1 v_2 - u_2 v_1)\hat{k}$$

$$= [(u_2 v_3 - u_3 v_2), (u_3 v_1 - u_1 v_3), (u_1 v_2 - u_2 v_1)]^T$$

HOMOGENEOUS CO-ORDINATES: 2D

Lines

Given two lines l_1 and l_2, what is the homogeneous co-ordinate vector p for the point of their intersection ?

Same idea: $l_1^T p = p^T l_1 = 0$ and $p^T l_2 = 0$

$$p = l_1 \times l_2$$

- Cross product between two points gives us the line between them
- Cross product between two lines gives us the point common to both
- What happens if l_1 and l_2 are parallel ?

Answer: Third co-ordinate of $l_1 \times l_2$ is 0. Point at infinity.

HOMOGENEOUS CO-ORDINATES: 2D

Transformations

- Translation:
  - $x' = x - c_x, y' = y - c_y$

$$p' = \begin{bmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{bmatrix} p$$

- Verify this works for any scaled version of T above
- Verify this works for $p = [\alpha x, \alpha y, \alpha]$, for any $\alpha \neq 0$

HOMOGENEOUS CO-ORDINATES: 2D

Transformations

- Rotation Around the Origin
  - $x' = x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta$

$$p' = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} p$$

- Translation around a different point c_x, c_y ?

$$p' = \begin{bmatrix} 1 & 0 & c_x \\ 0 & 1 & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -c_x \\ 0 & 1 & -c_y \\ 0 & 0 & 1 \end{bmatrix} p$$

HOMOGENEOUS CO-ORDINATES: 2D

Transformations

- Euclidean Transformation

$$p' = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} p$$

  - R is a 2×2 rotation matrix, $R^T R = I$
  - t is a 2×1 translation vector
  - 0^T here represents a 1×2 row of two zeros
  - Preserves orientation, lengths, areas

If $R^T R = I$, is R always of the form:

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

HOMOGENEOUS CO-ORDINATES: 2D

Transformations

- Euclidean Transformation

$$p' = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} p$$

  - R is a 2×2 rotation matrix, $R^T R = I$
  - t is a 2×1 translation vector
  - 0^T here represents a 1×2 row of two zeros
  - Preserves orientation, lengths, areas

If $R^T R = I$, is R always of the form:

$$R = \begin{bmatrix} -\cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

HOMOGENEOUS CO-ORDINATES: 2D

Transformations

- Euclidean Transformation

$$p' = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} p$$

  - R is a 2×2 rotation matrix, $R^T R = I$
  - t is a 2×1 translation vector
  - 0^T here represents a 1×2 row of two zeros
  - Preserves orientation, lengths, areas
- Isometries
  - $R^T R = I$ can also correspond to reflections

$$R = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

  - If we allow this in R above, more general than euclidean
  - Preserves lengths, areas, but not orientation.

HOMOGENEOUS CO-ORDINATES: 2D

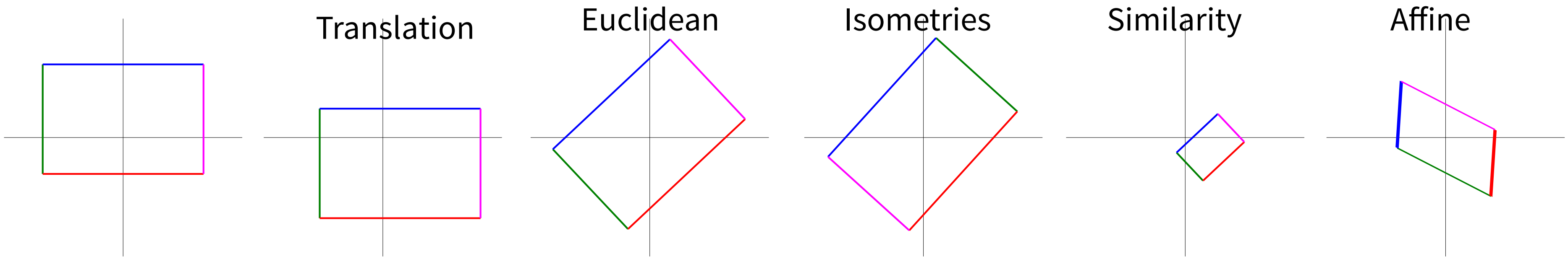
Transformations

What about scaling ?

Allow uniform scaling s along both co-ordinates:

$$p' = \begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix} p$$

Called a similarity: preserves ratio of lengths, angles.



HOMOGENEOUS CO-ORDINATES: 2D

Transformations

Affine Transformation

$$p' = \begin{bmatrix} A & t \\ 0^T & 1 \end{bmatrix} p$$

where A is a general invertible 2×2.

Preserves ratios of areas, parallel lines stay parallel.

Prove that parallel lines stay parallel.

- Consider the homogeneous vector q for intersection of two lines that are parallel.
- The third co-ordinate of q is 0, because the lines don't intersect.
- The affine transform doesn't change the third co-ordinate.
- Hence, the lines still intersect at infinity after the transformation.

HOMOGENEOUS CO-ORDINATES: 2D

Most general form:

$$p' = Hp$$

where H is a general invertible 3×3 matrix.

- Called a projective transform or **homography**.
- All bets are off! Parallel lines can now intersect. Maps quadrilaterals to quadrilaterals.
- Defined upto scale. So 8 degrees of freedom.
- Hierarchy of Transforms
  - Translation (2 dof) < Euclidean (3 dof) < Affine (6 dof) < Homography (8 dof)
- Defines mapping of co-ordinates of corresponding points in two images taken from different views:
  - If all corresponding points lie on a plane in the world.
  - If only the camera orientation has changed in two views (center is at the same place)

ESTIMATION

Homography

$\xrightarrow{\hspace{3em}}$

H

(0,0) (W,0)

(0,H) (W,H)

Say I knew the length and width of the rug.

Estimate H,
apply to all points,
measure lengths in meters !

I know a bunch of pairs of points (p_i', p_i), and want to find H such that:

$$p_i' \sim H p_i, \quad \forall i$$

ESTIMATION

I know a bunch of pairs of points (p_i', p_i), and want to find H such that:

$$p_i' \sim Hp_i, \quad \forall i$$

- Equality only upto scale: how do you turn that into an equation ?
- How many unknowns ? 8 (defined upto scale)
- How many equations for four points? 8 (2 x 4)

But how do we write these equations for equality upto scale ?

$$p_i' \times (Hp_i) = 0$$

Recall: $u \times v = [(u_2 v_3 - u_3 v_2), (u_3 v_1 - u_1 v_3), (u_1 v_2 - u_2 v_1)]^T$

ESTIMATION

$$p_i' \times (Hp_i) = 0$$

This is a linear equation in the elements of H.

Let $h = [h_1, h_2, h_3, h_4, h_5, \dots h_9]$ be a vector of the 9 elements of H. Can write:

$$A_i h = 0$$

What is the size of A_i ?

The cross product gives us 3 equations, so A_i is 3×9.

But, one of the rows of A_i is a linear combination of the other (A_i has rank 2). Can choose to keep only two rows, or all three.

Stacking all the A_i matrices for all different correspondences, we get:

$$Ah = 0$$

A is $2n \times 9$ or $3n \times 9$ matrix, where n is number of correspondences. Rank(A) is at most $2n$.

Rank exactly equal to $2n$ if no three points are collinear.

ESTIMATION

So we have $Ah = 0$ and want to find h upto scale. A has rank $2n$ and h has 9 elements.

Case 1: $n = 4$ non-collinear points.

- Trivial solution is $h = 0$. But want to avoid this.
- Cast as finding $Ah = 0$ such that $||h|| = 1$.
- Since A is exactly rank 8, there exists such a solution and it is unique (upto sign).
- Can find using eigen-decomposition / SVD.
- $A = UDV^T$ where D is diagonal with last element 0. h is the last column of V.

Case 2: $n > 4$ non-collinear points.

- Over-determined case. Want to find "best" solution.
- $h = \arg\min_h ||Ah||^2, \quad ||h|| = 1$
- Same solution, except that instead of taking 0 singular value, we take minimum singular value.
- $||Ah||^2 = (Ah)^T(Ah) = h^T(A^TA)h$
- Minimized by unit vector corresponding to lowest eigenvalue of A^TA, or lowest singular value of A.

ESTIMATION

Estimation from Lines

- How does a homography transform a line:

$$l^T p = 0 \leftrightarrow l'^T p' = 0$$

$$l^T H^{-1} H p = 0 \Rightarrow (H^{-T} l)^T (Hp) = 0$$

$$l' = H^{-T} l \Rightarrow l = H^T l'$$

- If we find four pairs of corresponding lines, we can get a similar set of equations for $l_i = H^T l_i'$ as for points.
- Get equations from $l_i \times (H^T l_i') = 0$for elements of H.

ESTIMATION

Other approaches:

- Instead of measuring $||Ah||^2$, might want to measure explicit geometric distance.
- Minimize distance in mapped cartesian co-ordinates (re-projection error).
- Involves division, no longer linear in H. Iterative methods.
- See "Multiple View Geometry in Computer Vision," Hartley & Zisserman: Section 4.2 (or really, the whole book for a thorough discussion of geometry)

NEXT TIME

- Estimate camera projection matrix
- Relate transformation between two views
- Automated matching to solve for depth: Stereo